МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ В. Н. КАРАЗІНА

Кафедра комп'ютерної фізики

"ЗАТВЕРДЖУЮ"

Директор навчально-наукового інституту

комп'ютерної фізики та енергетики

_____ Ірина ГАРЯЧЕВСЬКА

"30" 08 2024 р.

РОБОЧА ПРОГРАММА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ВСТУП ДО МЕТОДІВ МАШИННОГО НАВЧАННЯ

рівень вищої освіти	перший (бакалаврський)
галузь знань	10 Природничі науки
спеціальність	105 Прикладна фізика та наноматеріали
освітня програма	«Комп'ютерна фізика»
вид дисципліни	вибіркова

навчально – науковий інститут комп'ютерної фізики та енергетики

2024 / 2025 навчальний рік

Програму рекомендовано до затвердження вченою радою навчально-наукового інституту комп'ютерної фізики та енергетики

28 серпня 2024 року, протокол № 8/24

РОЗРОБНИКИ ПРОГРАМИ:
Нємченко Костянтин Едуардович, д.ф.-м.н., професор
Бакуменко Олександр Вікторович, старший викладач

Програму схвалено на засіданні кафедри комп'ютерної фізики

Протокол від 28 серпня 2024 року № 8/24

Завідувач кафедри комп'ютерної фізики

_____ Костянтин НЄМЧЕНКО

Програму погоджено з гарантом освітньо-наукової програми «Комп'ютерна фізика»

Гарант освітньо-наукової програми «Комп'ютерна фізика»

_____ Світлана РОГОВА

Програму погоджено науково-методичною комісією
навчально-наукового інституту комп'ютерної фізики та енергетики

Протокол від 28 серпня 2024 року № 8/24

Голова науково-методичної комісії
навчально-наукового інституту комп'ютерної фізики та енергетики

_____ Денис ПРОТЕКТОР

ВСТУП

Програма навчальної дисципліни «Вступ до методів машинного навчання» складена відповідно до освітньо-професійної програми «Комп'ютерна фізика» підготовки другого (магістерського) рівня вищої освіти

спеціальності (напряму) 105 Прикладна фізика і наноматеріали

1. Опис навчальної дисципліни

Предметом вивчення навчальної дисципліни є використання прикладних пакетів для задач фізики та енергетики.

1.1. Мета викладання навчальної дисципліни

Метою курсу «Вступ до методів машинного навчання» є вивчення та самостійне користування алгоритмами діагностики зображень або масивів даних з використанням самостійно розроблених алгоритмів та бібліотек для задач фізики та математики.

1.2. Основні завдання вивчення дисципліни

Основним завданням курсу «Вступ до методів машинного навчання» є застосування математичних алгоритмів для діагностування та керування складними системами керування процесами в фізиці та енергетиці.

У результаті вивчення даного курсу студент повинен

знати: методи діагностики даних за допомогою математичних алгоритмів та самостійне створення програмних продуктів, що реалізують алгоритми автоматизованої діагностики для певних задач в галузі фізики та енергетики.

вміти: застосовувати отримані знання на практиці при діагностиці та аналізу масивів даних за допомогою методів паралельних обчислень, узагальнювати вивчені алгоритмі на складні системи.

Для вивчення курсу необхідні знання з програмування, теорії алгоритмів, математичного аналізу, диференціальних рівнянь, математичної фізики, та курсів лінійної алгебри та аналітичної геометрії.

Основними формами викладання навчального матеріалу з дисципліни «Прикладне машинного навчання» є лекції, лабораторні заняття та самостійна робота студентів.

Методика викладання ґрунтується на проведенні лекційних, та лабораторних у 2-му семестрі, із загальним аудиторним обсягом 64 годин, що разом з 56 годинами самостійної роботи складає 120 годин загального обсягу. Лекційні заняття проводяться методом семінару. Лабораторні заняття – шляхом розробки окремими студентами завдань програм на мові Python в певному середовищі розробки за вибором студента.

1.3. Кількість кредитів 4

1.4. Загальна кількість годин 120

1.5. Характеристика навчальної дисципліни	
Вибіркова	
Денна форма навчання	Заочна (дистанційна) форма навчання
Рік підготовки	
1-й	-й
Семестр	
2-й	-й
Лекції	
32 год.	год.
Практичні, семінарські заняття	
0 год.	год.
Лабораторні заняття	
32 год.	год.
Самостійна робота	
56 год.	год.
Індивідуальні завдання	
год.	

1.6. Заплановані результати навчання

В результаті вивчення навчальної дисципліни студенти оволодіють засобами застосування математичних алгоритмів для діагностування та керування складними системами керування процесами в фізиці та енергетиці.

1. ПРН1. Використовувати знання в галузі прикладної фізики, математики, електроніки та інформаційних технологій для виконання наукових досліджень та розв’язання виробничих задач.
2. ПРН2. Знаходити та аналізувати наукову та науково-технічну інформацію в галузі прикладної фізики та наноматеріалів із вітчизняних та зарубіжних джерел, в тому числі з використанням сучасних пошукових систем.
3. ПРН4. Встановлювати та аргументувати нові залежності між параметрами та характеристиками фізичних систем.

2. Тематичний план навчальної дисципліни

Розділ 1. Мова програмування Python

Тема 1. Основні галузі використання

Опис різних сфер, де застосовуються програмні технології — від веб-розробки до наукових досліджень і автоматизації промислових процесів.

Тема 2. Основні правила програмування

Вивчення принципів написання чистого, ефективного та читабельного коду, таких як уникання дублювання, використання коментарів та зрозумілих імен змінних.

Тема 3. Основні оператори

: Операції над даними, включаючи арифметичні, логічні, порівняння, присвоєння та інші основні оператори, що використовуються в мовах програмування.

Тема 4. Структури даних Вивчення різних типів структур даних, таких як масиви, списки, дерева, графи, стеки та черги, які допомагають організовувати та обробляти дані.

Тема 5. Введення та виведення даних

Операції з введенням та виведенням інформації з програми, робота з файлами, консолями та іншими потоками вводу/виводу

Тема 6. ООП

Принципи ООП, такі як інкапсуляція, наслідування, поліморфізм, що дозволяють створювати модульний та масштабований код.

Тема 7. Стандартна бібліотека

Огляд бібліотек, які постачають основні функції для роботи з рядками, колекціями, файлами, математичними операціями та ін.

Тема 8. Особливості використання функцій

Як створювати та використовувати функції в програмуванні, їх параметри, повернення значень та області видимості.

Розділ 2. Загальна теорія розпізнавання

Тема 1. Задача розпізнавання

Вивчення основ задач розпізнавання, зокрема, як навчити комп'ютер ідентифікувати об'єкти або образи в даних на основі певних ознак чи характеристик.

Тема 2. Математична постановка задачі класифікації. Способи визначення критеріїв відмінності

Розгляд теоретичних основ класифікації, побудова математичних моделей для розподілу об'єктів за класами, а також критерії, що дозволяють відрізняти ці класи.

Тема 3. Класифікація методів розпізнавання. Паралельна та послідовна класифікації. Створення критеріїв розпізнавання

Ознайомлення з різними методами розпізнавання, такими як статистичні, нейронні мережі, машинне навчання, а також паралельні та послідовні підходи до класифікації.

Тема 4.Основні алгоритми автоматизованої діагностики

Вивчення алгоритмів, що використовуються для автоматичної діагностики, таких як методи машинного навчання, які дозволяють ідентифікувати або діагностувати проблеми в системах на основі вхідних даних.

Тема 5. Критерії вибору алгоритмів. Алгоритми власних значень та матриць.

Розгляд критеріїв, за якими вибираються алгоритми для конкретних задач, зокрема, алгоритмів для роботи з власними значеннями та матрицями, які використовуються у методах розпізнавання.

Тема 6. Алгоритми створення моделей

Вивчення алгоритмів, що використовуються для створення математичних та статистичних моделей, які можуть бути використані в задачах розпізнавання.

Тема 7. Розпізнавання зображень.

Поглиблене вивчення методів обробки та класифікації зображень, використовуючи методи комп'ютерного зору та глибинного навчання для розпізнавання об'єктів, сцен чи патернів на зображеннях.

Тема 8 Розпізнавання символів. Розпізнавання рукописних текстів

Тема, що охоплює методи розпізнавання тексту, зокрема, алгоритми для автоматичного розпізнавання друкованих і рукописних символів, а також їх практичне застосування для перетворення рукописного тексту в цифровий формат.

Пояснення та приклади за темами

Тема 1. Задача розпізнавання

Пояснення: Задача розпізнавання полягає в тому, щоб навчити машину визначати або ідентифікувати об'єкти, символи або патерни на основі наданих даних. Це може бути як розпізнавання образів, так і аналіз текстів або голосових команд.

Приклад: Якщо система повинна розпізнавати зображення котів, то алгоритм аналізує різні характеристики зображення, такі як контури, текстури, кольори, для визначення, чи є на ньому кіт чи ні.

Тема 2. Математична постановка задачі класифікації. Способи визначення критеріїв відмінності

Пояснення: Класифікація — це процес сортування об'єктів за різними категоріями чи класами. Математично задача класифікації полягає в пошуку функції, яка на основі вхідних даних визначає, до якого класу належить об'єкт.

Приклад: У задачі класифікації, де потрібно розпізнати види квітів (тюльпан, троянда, лілія), критерії відмінності можуть бути, наприклад, форма пелюсток, колір, висота рослини. Всі ці характеристики визначають клас, до якого належить квітка.

Тема 3. Класифікація методів розпізнавання. Паралельна та послідовна класифікація

Пояснення:

- **Паралельна класифікація** передбачає, що дані одночасно обробляються кількома різними алгоритмами (наприклад, кожен клас має свій окремий класифікатор).
- **Послідовна класифікація** передбачає, що один алгоритм класифікує об'єкти за одним критерієм, потім інший — за іншим, і так далі.

Приклад:

- **Паралельна класифікація**: кілька моделей одночасно класифікують об'єкти на "кішки" або "собаки".
- **Послідовна класифікація**: спочатку визначається, чи є об'єкт тварина, потім, якщо так, класифікується як "кішка" чи "собака".

Тема 4. Основні алгоритми автоматизованої діагностики

Пояснення: Алгоритми автоматизованої діагностики використовуються для визначення, чи є у системи проблеми, а також для їх класифікації. Це можуть бути алгоритми машинного навчання, які використовують дані для побудови моделей і подальшої їх діагностики.

Приклад: Алгоритм може аналізувати медичні зображення (наприклад, рентгенівські знімки) і автоматично виявляти можливі ознаки хвороби (наприклад, рак легенів).

Тема 5. Критерії вибору алгоритмів. Алгоритми власних значень та матриць

Пояснення: При виборі алгоритму важливо враховувати характеристики задачі: точність, швидкість обробки, вимоги до ресурсів. Алгоритми власних значень використовуються для зниження вимірності даних, наприклад, у методах головних компонент (PCA).

Приклад: У задачі розпізнавання зображень алгоритм може вибирати важливі ознаки (власні компоненти) для зменшення розміру вхідних даних без втрати значної інформації.

Тема 6. Алгоритми створення моделей

Пояснення: Моделі створюються на основі даних для вирішення конкретних задач. Це можуть бути як статистичні, так і машинно-навчені моделі.

Приклад: Алгоритм на основі навчання на прикладах будує модель, яка може передбачати результат для нових невідомих даних (наприклад, модель для прогнозування ціни на нерухомість).

Тема 7. Розпізнавання зображень

Пояснення: Розпізнавання зображень полягає в тому, щоб на основі візуальних даних (пікселів, контурів, текстур) визначити, що зображено на фото.

Приклад: Використання нейронних мереж для автоматичного визначення предметів на зображеннях (наприклад, розпізнавання автомобіля, людини, дерева).

Тема 8. Розпізнавання символів. Розпізнавання рукописних текстів

Пояснення: Розпізнавання символів включає ідентифікацію букв, цифр або інших знаків на зображеннях. Розпізнавання рукописного тексту — це особливий випадок, оскільки рукописний текст може мати багато варіантів написання.

Приклад: Алгоритм OCR (оптичне розпізнавання символів) може використовуватись для сканування рукописних документів і перетворення їх у текстовий формат.

Приклади практичних завдань:

1. **Задача: Розпізнавання образів** Скачайте набір даних з відомим набором образів, наприклад, "CIFAR-10" або "MNIST". Використовуючи бібліотеку **TensorFlow** або **PyTorch**,

створіть просту модель, яка буде класифікувати зображення за категоріями (наприклад, цифри або об'єкти).

2. **Задача: Класифікація даних** Використовуйте набір даних з відкритих джерел (наприклад, **Iris dataset**) для побудови моделі класифікації з використанням алгоритму K-ближчих сусідів або дерева рішень.

3. **Задача: Розпізнавання тексту (OCR)** Створіть простий проект для розпізнавання тексту на зображеннях за допомогою бібліотеки **Tesseract OCR**.

3. Структура навчальної дисципліни

Назви розділів	Кількість годин											
	денна форма						заочна форма					
	усього	у тому числі					усього	у тому числі				
		л	п	лаб.	інд.	с. р.		л	п	лаб.	інд.	с. р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Розділ 1. Стандартні бібліотеки шаблонів												
Тема 1	6	2		2		2						
Тема 2	6	2		2		2						
Тема 3	6	2		2		2						
Тема 4	6	2		2		2						
Тема 5	8	2		2		4						
Тема 6	8	2		2		4						
Тема 7	8	2		2		4						
Тема 8	8	2		2		4						
Разом за розділом 1	56	16		16		24						
Розділ 2. Паралельні обчислення												
Тема 1	8	2		2		4						
Тема 2	8	2		2		4						
Тема 3	8	2		2		4						
Тема 4	8	2		2		4						
Тема 5	8	2		2		4						
Тема 6	8	2		2		4						
Тема 7	8	2		2		4						
Тема 8	8	2		2		4						
Разом за розділом 2	64	16		16		32						
Усього годин	120	32		32		56						

4. Теми практичних занять

Не передбачені навчальним планом

Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1	Основні галузі використання	2
2	Основні правила програмування	2
3	Основні оператори	2
4	Структури даних	2
5	Введення та виведення даних	2
6	ООП	2
7	Стандартна бібліотека	2
8	Особливості використання функцій	2

9	Робота з зображеннями на мові Python	2
10	Використання мови Python в системах керування	2
11	Задача розпізнавання	2
12	Способи визначення критеріїв відмінності	2
13	Створення критеріїв розпізнавання	2
14	Основні алгоритми автоматизованої діагностики	2
15	Алгоритми власних значень та матриць.	2
16	Алгоритми створення моделей	2
	Разом	32

5. Завдання для самостійної робота

№ з/п	Назва теми	Кількість годин
1	Вивчити основні галузі використання	2
2	Згадати оосновні правила програмування	2
3	Засвоїти основні оператори	2
4	Використати основні типи структур даних	2
5	Вивчити введення та виведення даних для задач МН	4
6	Засвоїти основні методи ООП	4
7	Використати стандартні бібліотека	4
8	Вивчити особливості використання функцій	4
9	Провести робота з зображеннями на мові Python	4
10	Ознайомитись з використанням мови Python в системах керування	4
11	Визначити задача розпізнавання	4
12	Вивчити способи визначення критеріїв відмінності	4
13	Засвоїти створення критеріїв розпізнавання	4
14	Використати в конкретних задачах певні алгоритми автоматизованої діагностики	4
15	Вивчити алгоритми власних значень та матриць.	4
16	Ознайомитись з алгоритмами створення моделей	4
	Разом	56

6. Індивідуальні завдання

7. Методи контролю

Навчальна програма нормативної дисципліни «Вступ до методів машинного навчання» відповідає чинним нормативним документам, рекомендованим Міністерством освіти і науки України як навчальна програма для студентів фізичного спрямування для вищих навчальних закладів. Нормативна навчальна дисципліна «Вступ до методів машинного навчання» є складовою циклу професійної підготовки фахівців освітньо–кваліфікаційного рівня «магістр».

Зміст програми, об’єм учбових питань дисципліни «Вступ до методів машинного навчання» визначаються потребою загальнонаукової, загально-інженерної та технічної підготовки.

Послідовність викладання тем не є суворо обов’язковою і може бути змінена у відповідності із структурно-логічною схемою викладання інших дисциплін.

Вивчення дисципліни «Вступ до методів машинного навчання» передбачає викладання лекцій, проведення практичних занять, виконання лабораторних робіт. Підсумковий контроль знань здійснюється на екзамені.

На лекціях викладається теоретичний матеріал, який ілюструється типовими прикладами і задачами за профілем підготовки фахівців з прикладної фізики та наноматеріалів. Викладання лекційного матеріалу повинно має закінчений характер, здійснюється у доступній і наочній формі, містить проблемні ситуації.

Основною метою практичних занять є розвиток навичок практичного застосування вищої математики і закріплення теоретичного матеріалу. При вирішенні задач рекомендується користуватися стандартними прийомами і методиками. На кожному практичному занятті частину учбового часу доцільно використовувати для самостійного розв'язку задач, контролюючи при цьому правильність обраного методу і ходу рішення.

Важливим фактором засвоєння курсу «Вступ до методів машинного навчання» й оволодіння її методами є самостійна робота студентів. Для самостійного відпрацювання розділів і тем дисципліни пропонуються лабораторні роботи, при перевірці виконання яких здійснюється ефективний контроль за рівнем засвоєння матеріалу.

На всіх заняттях підкреслюється роль і значення предмету на конкретних прикладах і задачах прикладного характеру, висвітлюється її зв'язок з іншими дисциплінами і майбутньою професією.

8. Схема нарахування балів

Поточний контроль, самостійна робота, індивідуальні завдання	
Розділ 1	
Т1	2
Т2	2
Т3	2
Т4	2
Т5	2
Т6	2
Т7	2
Т8	2
РГР	14
Розділ 2	
Т1	2
Т2	2
Т3	2
Т4	2
Т5	2
Т6	2
Т7	2
Т8	2
РГР	14
Залік	40
Сума	100

Т1, Т2 … Т8 – теми розділів.

Мінімальна кількість балів для допуску студента до підсумкового заліку становить – 15 балів.

Критерії оцінювання

З кожну тему - 2 бали дається одне завдання (задача), розв'язок якої оцінюється за такими критеріями:

Наявність відповіді – 0.25 бал
Коректність викладок – 0.25 бал
Логічність викладок – 0.5 бал
Коректність визначень – 0.5 бал
Повнота відповіді – 0.25 бал
Знання цілей задачі – 0.25 бал

Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка	
	для екзамену	для заліку
90 – 100	відмінно	зараховано
70-89	добре	
50-69	задовільно	
1-49	незадовільно	не зараховано

9. Рекомендована література

Основна література

1. Ефективне використання STL Скотт Мейєрс 4.5 ISBN: 5-94723-382-7 2002
2. C++ та STL. Девід Р. Мюссер, Жилмер Дж. Дердж, Атул Сейні 5.0 ISBN: 978-5-8459-1665-5, 978-0-321-70212-8: 2010
3. Розширення бібліотеки STL для C++. Набори та ітератори (+ CD… Метью Вілсон 0.0 ISBN: 978-5-94074-442-9, 978-5-9775-0196-5, 978-0-321-30550-7 2008
4. Узагальнене програмування та STL. Використання та нарощування… Метью Г. Остерн 5.0 ISBN: 5-7940-0119-4, 0-201-30956-4 2004

Допоміжна література

1. STL. Кишеньковий довідник Рей Лішнер 0.0 ISBN: 5-469-00389-2, 0596005563 2005
2. Стандартна бібліотека C++. Ніколаї М. Джосаттіс 4.5 ISBN: 978-5-8459-1837-6 2014

10. Посиланная на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. http://www.rsdn.ru/res/book/cpp/josattis.xml
2. http://www.rsdn.ru/res/book/cpp/halpern.xml